Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 8 (2012) 345 – 350

Procedia Computer Science

New Challenges in Systems Engineering and Architecting
Conference on Systems Engineering Research (CSER)
2012 – St. Louis, MO
Cihan H. Dagli, Editor in Chief
Organized by Missouri University of Science and Technology

System for Detection of Malicious Wireless Device Patterns

Shikhar P Acharya[a], Ritesh Arora[b], Ivan G. Guardiola[c],a*

[a,b,c]*Missouri University of Science and Technology,1870 Miner Circle, Rolla, MO, 65409*

Abstract

The research within presents the use of Hidden Markov Models (HMM) for the detection of wireless devices in highly noisy environments using their unintended electromagnetic emissions (UEE). All electromagnetic devices emit such radiation that is unique to the electronics, housing, and other device attributes. This pattern recognition system can provide continuous detection analysis and can provide ideal information regarding the distance to an unknown device. An experiment was performed where UEE of a device was detected by a spectrum analyzer. Experimental result shows that our model can accurately detect if there is a device nearby emitting UEE or not.

Keywords: malicious device identification, hidden markov model, unintended electromagnetic emissions

1. Introduction

Improvised explosive devices (IEDs) are extensively used by terrorist to attack their enemies. Radio frequency (RF) receivers are primarily used as detonators of these IEDs. There are two ways RF receivers can be used as triggers. One way is to wire the explosive material to the RF receiver directly. A call is then made to that phone which results in the generation of a small charge. This causes the trigger to the wired IED which results in explosion. This method is called 'call in' technique. Another technique is to set a timer using the internal alarm of the RF receiver. The IED will detonate when the alarm is triggered [1].

A large number of coalition fatalities in Afghanistan and Iraq are due to the IEDs. Table 1 illustrates the number of coalition fatalities in Afghanistan and the number and percentage of fatalities attributed to

* Corresponding author. Tel.: +1-573-341-6556
address: spa2p7@mst.edu

1877-0509 © 2012 Published by Elsevier B.V. Open access under CC BY-NC-ND license.
doi:10.1016/j.procs.2012.01.071

IEDs. If we consider the total coalition fatalities of 2,683 in and around Afghanistan, IEDs are responsible for 42% of deaths [2]. Refer to table 1 for data. If we just consider 2,156 hostile fatalities of the coalition force, then IEDs are responsible for a staggering 52% of all fatalities. If we can somehow detect IEDs that are present nearby, the number of coalition death can be significantly reduced not only in and around Afghanistan, but in other hostile environment where this form of terrorism exists. This shows that detection and localization of RF receivers in a hostile territory could be an effective approach to reduce the number of fatalities.

Year	Coalition Fatalities in Afghanistan	Fatalities Attributed to IED	Percentage Attributed to IED
2001	12	0	0%
2002	70	4	6%
2003	58	3	5%
2004	60	12	20%
2005	131	20	15%
2006	191	41	21%
2007	232	78	34%
2008	295	152	52%
2009	521	275	53%
2010	711	368	52%
2011	402	177	44%
Total	2683	1130	42%

Table 1: Coalition Fatalities in Afghanistan

All RF receivers produce unintended electromagnetic emissions (UEE) [3]. RF receivers are based on superheterodyne receiver architecture of Edwin Armstrong. This architecture uses local oscillator (LO) as a necessary component of the receiver. Some of the electromagnetic emission inevitably leaks from the LO and is emitted from the antennae of the RF receivers as UEE. This process is explained in [4]. The emissions also depend on specific electronics and the housing of the device. If we identify UEEs of possible explosives in hostile area, IEDs can be safely deactivated prior to causing harm.

Unfortunately, UEEs are very weak in power and is thus difficult to detect, at least from a significant distance. The challenge is to identify such signal in the presence of noise. Some progress has been done in this area. Neural Network has successfully been employed to identify and locate electronic devices up to a distance of 10m [5]. Dong *et al.* (2006) have successfully identified three vehicles based on their RF emissions using neural networks [6]. But as noted by them, the problem with these approaches is that all these methods are sensitive to noise. A noble approach of using statistical correlation to identify unintended emissions from a toy truck is given in [7]. A biconical antenna was used and the measurements were taken by oscilloscope for time domain and spectrum analyzer for frequency domain. Ideal unintended electromagnetic emission pulse was constructed from the observations using the cascading correlation procedure. The cascaded signal is finally normalized by the maximum value to get

an ideal pulse. The detection is done by correlating the test signal with ideal pulse. A threshold correlation value is identified above which the signals are assumed from the particular device. This method can identify devices up to a distance of 40m with an accuracy of 98% under Receiver Operating Characteristics (ROC) curve. The drawback of this method is that, UEE changes with respect to battery charge, climate, and ambient noise. A little change in any of these parameters, which is very much likely in the real world, significantly reduces this methods' ability to correctly detect the IED.

We present a more robust approach to generate an underlying pattern recognition statistical model using HMM. This model once trained from emission data will outline a comparison between the trained model and the observations, which results in the highest likelihood estimation that a given observation came from a particular model. The reason for choosing HMM is that it is a robust model based upon well defined mathematical theories and which has proven results from a variety of applications that it performs well [8],[9].

2. Data Collection

Data collection for this project has been done using U3700 spectrum analyzer by Advantest and VERT400 tri-band antenna. XR150U Business Two-Way Radio is the RF receiver selected for this project. The operating frequency of the device is 450 – 470 MHz's. UEEs have very low power emissions. For the purpose of clarity, a span has been fixed to 20 kHz and the distance from which readings have been taken is 5 ft. Experimental setup for the data correction part is shown in figure 1. As we can see the RF receiver is kept at the fixed distance from spectrum analyzer. In order to collect data, the spectrum analyzer has been set to -30db level to appropriately display the signal on screen. On every instance of unintended emissions, one sample of 1,001 points are stored on excel sheet. 60 samples of data were collected altogether, 30 for device and 30 for noise.

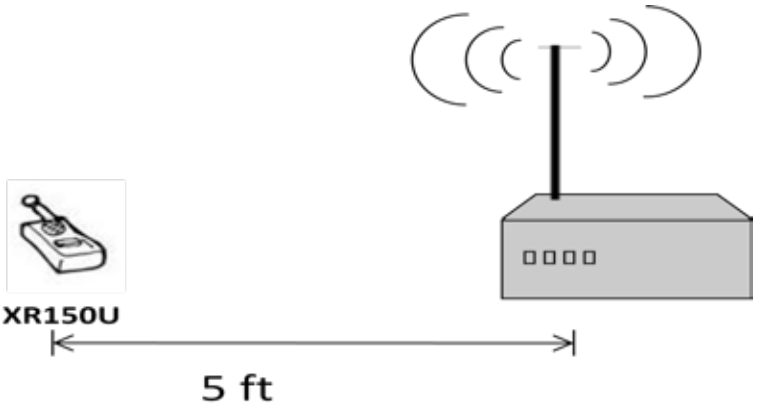

Figure 1: Experimental Setup for Data Collection

3. Data Preparation

The signal recorded by the spectrum analyzer constitutes the plot of amplitude against frequency. The reading was taken such that the central frequency always lies in the center of the span. The span of the signal is of 20 KHz. Each successive observation has amplitude measured at the difference of 20Hz. A typical signal is shown in figure 2a. Since the span is chosen such that the effect of UEE is visible around

the center frequency, the observation near the center frequency is of importance to us. We considered 401 observations near the center and crop the rest of 600 observations. We limit ourselves with the observation between 300 and 700. The resulting figure is shown in figure 2b.

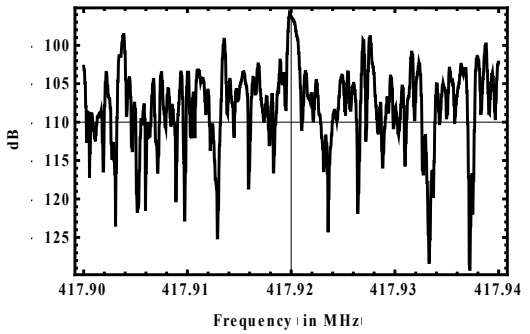

Figure 2a: Electromagnetic Emission with 1001 Observations

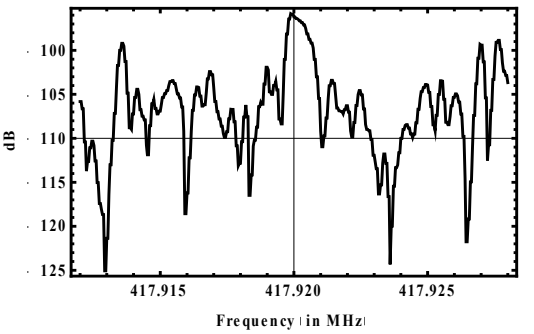

Figure 2b: Electromagnetic Emission with 401 Observations

4. Methodology

Hidden markov model is represented by the tuple $\lambda = (M,N,A,B, \pi)$ where each element is defined as following:

i. M is the number of states and individual state is represented as $S = \{S_1, S_2,\ldots,S_m\}$
ii. N is the number of observations and individual observation is represented as $O = \{O_1, O_2,\ldots,O_n\}$
iii. $A = \{a_{ij}\}$ is the state transition matrix where $a_{ij} = P[q_{t+1} = S_j, / \quad q_t = S_i]$, $1 \leq i, j \leq M$
iv. $B = \{b_j(k)\}$ is the emission transition matrix where $b_j(k) = P[O_k \text{ at } t / q_t = S_j]$, $1 \leq i \leq M$ and $1 \leq k \leq N$
v. $\pi = \{ \pi_i\}$ is the initial state probability distribution where $\pi_i = P[q_1 = S_i]$, $1 \leq i \leq M$

To represent any system as HMM described above, we need to have a Markov Chain and for this, we should define some states such that there is some probability associated with transfer from one state to another as stated in *iii* above. The change in amplitude of successive observation is taken as a state. The observations are observed at a span of 20 Hz. We have classified the difference in successive observations into five classes and these are taken as the 5 states that define HMM. The state definition is illustrated in table 2a. Since HMM is a doubly embedded stochastic process, we need to define another stochastic process apart from the markov chain. The set of visible observation sequence completes the components required to define the HMM along with the markov chain. We have defined the amplitude of the signal corresponding to the change in observation defined in state as the observation for our HMM. We have classified the amplitude into six classes and these are taken as the 6 observations for our HMM. The probability of observing one of the six observations defined in table 2b from a state gives the Emission Probability Matrix (EPM). A typical HMM is represented in figure 3. Using the principle of counting, we calculate the probability of observing a particular observation coming from a given state. Wolfram Mathematica 7 was used to calculate the Transition Probability Matrix (TPM) and EPM as defined in *iii* and *iv* respectively.

State Name	*Power change range (in dB)*
State 1	> 4dB
State 2	(0, 4] dB
State 3	0 dB
State 4	[-4,0) dB
State 5	< -4dB

Table 2a: State Definition

Observation	*Amplitude level range (in dB)*
Observation 1	< -120dB
Observation 2	[-120, -115)dB
Observation 3	[-115, -110)dB
Observation 4	[-110, -105)dB
Observation 5	[-105, -100)dB
Observation 6	>= -100 dB

Table 2b: Observation Definition

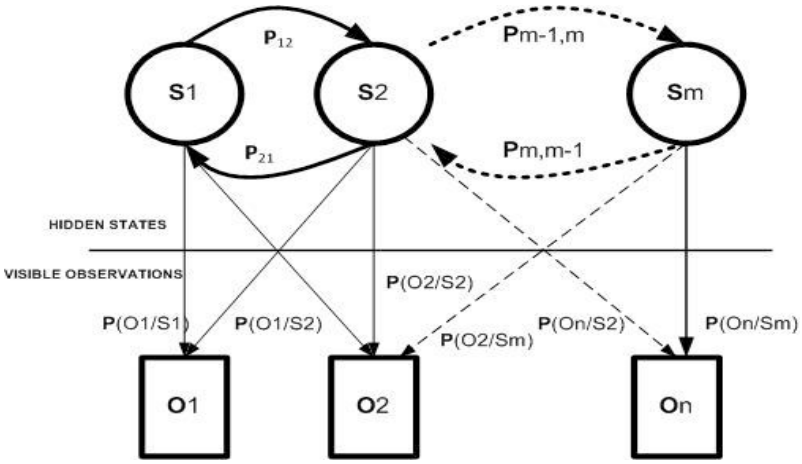

Figure 3: HMM Representation

5. Experimental Results

We have two hidden markov models, one defined for the device and another defined for noise. A total of 20 data samples, 10 from device and 10 from noise were fed as inputs to both models. The probability that the model generated the sample data was calculated using Matlab 7.10. The corresponding device or noise of the model that generated higher probability was selected as the one generating the signal. For example if P ($O_1/\boldsymbol{\lambda d}$) > P ($O_1/\boldsymbol{\lambda n}$) where O_1 is an observation signal, $\boldsymbol{\lambda d}$ is HMM for the device and $\boldsymbol{\lambda n}$ is HMM for the noise, then we infer that O_1 is emitted by $\boldsymbol{\lambda d}$. Our model was accurately able to identify the source of all the input signals: all the 10 UEE signals emitted from the device were identified as device signals and the rest of noise signals were accurately identified as noise. This is illustrated in table 3.

	Noise Signal	*Device Signal*
Number of times identified as device	0	10
Number of times identified as noise	10	0

Table 3: Experimental Results

6. Conclusions and Future Works

We have successfully employed HMMs to detect UEE. As per our knowledge, this is the first time HMMs have been applied in identifying malicious devices through their unintended emissions. We have proved that at a shorter distance, we can infer whether there is a malicious device or not even in a noisy environment without using an amplifier. This research has provided a foundation for future work. The first exploration to be done is to verify if this method works on longer distances than is considered here. Another exploration would be to determine if this method can recognize between two or more RF receivers instead of just differentiating between a single device and noise. Finally, another future direction would be to build a model with various features other than taking the difference in amplitude of successive observations.

References

[1] T. Strother, "Cell Phone Use by Insurgents in Iraq", Urban Warfare Analysis Center, 2007

[2] Retreived from http://icasualties.org/oef/ on 06/09/2011

[3] B. Wild and K. Ramchandran, "Detecting Primary Receivers for Cognitive Radio Applications", in Proc. of the First IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks, 2005

[4] S. M. Weiss, R. D. Weller., and S. D. Driscoll. "New measurements and predictions of UHF television receiver local oscillator radiation interference". Available: h-e.com/pdfs/rw_bts03.pdf

[5] H. Weng, X. Dong, X. Hu, D. G. Beetner, T. Hubing and D. Wunsch, "Neural Network Detection and Identification of Electronic Devices Based on Their Unintended Emissions", IEEE Transactions on Electromagnetic Compatibility, 2006

[6] X. Dong, H. Weng, D. G. Beetner, T. H. Hubing, D. C. Wunsch, M. Noll, et al, "Detection and Identification of Vehicles Based on Their Unintended Electromagnetic Emissions", IEEE Transactions on Electromagnetic Compatibility, 2006

[7] S. A. Seguin, D. Beetner, T. Hubing "Detection and Identification of Low-Cost RF Receivers Based on their Unintended Electromagnetic Emissions", submitted to IEEE Transactions on Electromagnetic Compatibility, 2009

[8] Rabiner, "A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", Proceedings of the IEEE, 1989

[9] S. R. Eddy, "What is Hidden Markov Model", Nature Biotechnology, 2004